REMARKS

Claims 38-57 are pending in the application. Claims 1 and 6-8 were rejected under 35 U.S.C. § 112, second paragraph, as described in paragraph 2 of the Office Action. Claims 1-7, 9-15, 22-34 and 37 were rejected under 35 U.S.C. § 103 as described in paragraph 4 of the Office Action. Claims 8, 17-18 and 35-36 were rejected under 35 U.S.C. § 103 as described in paragraph 5 of the Office Action. Claims 19 and 20 were rejected under 35 U.S.C. § 103 as described in paragraph 6 of the Office Action. Claim 21 was rejected under 35 U.S.C. § 103 as described in paragraph 7 of the Office Action. Claims 38, 44 and 51-57 are the only independent claims.

The specification has been amended to correct typographical errors and to generally place the application in correct idiomatic English. Attached hereto is a marked-up version of the changes made to the specification by the current amendment. The attachment is captioned "**Version with Markings to Show Changes Made**."

Additionally attached hereto are Replacement Formal Drawings for Figs. 2, and 5-8. The "audio-scene discrimination means" has been changed to --audio-scene identification means-- in item 10 of Fig. 2 to correspond with the disclosure and the "sream" has been changed to --stream-- in item 71 of Fig. 5, item 81 of Fig. 6, item 91 of Fig. 7 and item 110 in Fig. 8.

It is respectfully submitted that the outstanding rejections of the claims are moot, as the claims have been cancelled.

It is respectfully submitted that claims 38-57 have been drafted in compliance with 35 U.S.C. § 112.

It is respectfully submitted that claims 38-57 are patentable over the prior art of record for the following reasons.

The present invention relates to controlling a dynamic virtual space represented by 3-dimensional computer graphics (CGs), static images, dynamic images, audio and text which are based on a network such as the Internet.

There are demands in the market for enabling real-time CGs operation by a user from an information terminal reproducing the CGs such as game machines and cellular phones.

The present invention enables a user to change motions of objects or parts of objects of a received CG stream. In particular, the present invention modifies the motion data, of an object to be

changed by the user, of the CG stream in accordance with operations by the user inputted through a user interface at a terminal side while reproducing the CG streams. The present invention enables a controlling of the CG objects in the environment as well as portions of CG objects from a terminal at the user's pleasure. The present invention additionally increases the degree of freedom in operations at the terminals that receive the CG streams. Still further, the present invention permits a user to: select a type and a number of objects of the CG stream that are received at the terminal side based on the user's preferences and operation abilities; and operate the selected objects, thereby increasing operating convenience. Finally, the present invention prevents the CG stream structure from breaking and permits common use of a reproduction means for reproducing the CG stream so that the CG stream can be provided to different terminals. These aspects are accomplished in part by replacing the motion data of the selected object within the CG stream with content inputted by the user.

Newly added independent claims 38 and 44 are drawn to a stream correction apparatus. Claim 38 requires, *inter alia*, a correction unit operable to generate a corrected stream by replacing the motion data of the selected component with data based on the operational contents inputted by the user interface unit and to output the corrected stream. Similarly, claim 44 requires, *inter alia*, a correction unit operable to generate a corrected stream by replacing the motion data of the selected object or object part with data based on the operational contents inputted by the user interface unit and to output the corrected stream.

Newly added independent claim 51 is drawn to a transmission and reception system. Claim 51 requires, *inter alia*, a correction unit that is operable to generate a corrected stream by replacing the motion data of the selected component with data based on the operational contents inputted by the user interface unit and to output the corrected stream.

Newly added independent claims 52-54 are drawn to a stream correction method, a computer graphics reproduction method and a computer graphics display method, respectively. Each of newly added independent claims 52-54 require, *inter alia*, correcting the input stream by replacing the motion data of the selected component with data based on the inputted operational contents.

Newly added independent claims 55-57 are drawn to a data storage medium having computer readable instructions stored thereon. Each of independent claims 55-57 require the computer

readable instructions capable of instructing a computer to, *inter alia*, correct the input stream by replacing the motion data of the selected component with data based on the inputted operational contents.

It is respectfully submitted that the applied prior art, either singly or in combination, fails to teach or suggest the above-identified limitations.

Matsuba discloses that motion information and audio information of CG characters can be manually controlled in real time and are sent to a plurality of terminals as a stream via a relay server. In accordance with Matsuba, at the plurality of terminals that receive the information, the CG characteristics and the shape data of the objects in the environment have been previously downloaded. Further, the shape data that has been previously downloaded and the operation information and the audio information of the CG characteristics received are drawn and reproduced using 3DCG browser software provided at the terminal. Accordingly, in Matsuba, the user can observe the CG animation from a favorable viewpoint.

According to Matsuba, users at a plurality of terminals can observe the CG animation in accordance with the received information using 3DCG software. However, Matsuba does not teach or suggest that users at the plurality of terminals can control the CG objects in the environment from their own terminal. More specifically, Matsuba does not disclose or suggest that the CG objects in the environment can be controlled by a user in accordance with that of the present invention. More importantly, though, Matsuba does not disclose or suggest correcting an input stream by replacing motion data of a selected component with data based on input operational contents in accordance with claims 38, 44 and 51-57. For this reason, Matsuba fails to teach that which is required in each of independent claims 38, 44 and 51-57.

Carson fails to teach the shortcomings of Matsuba such that a combination of the teachings of Matsuba in view of Carson would teach that which is required in each of independent claims 38, 44 and 51-57.

Carson relates to authoring shared virtual spaces (synchronous change of copies of virtual space data possessed by the respective clients), and a system for executing the authoring. Carson teaches using an off-the-shelf player and browser, for example Cosmo Player and Netscape, to arrive at a hybrid type of central control and distributed control for sharing objects (including actions of

avatars, and behaviors of shared objects). The system of Carson includes a gatekeeper, browser client, VRML extensions, a pre-processor and a wire protocol. The gatekeeper grants access to the shared virtual space, tracks ownership of shared objects, tracks and records the state of the virtual space, checks whether the participants are still connected to the virtual space and evicts those who have lost their network connections. The browser client is a Java applet running inside the web browser that is supplied with access to the shared virtual space. The browser client monitors events that occur in a copy of the VRML via the external authorizing interface and then multi casting these events to other clients. Further, the browser client receives notifications of events from other clients and then reflects the events to the virtual spaces held by the clients which received the notification. Still further, the browser client is customized by the pre-processor program. Duties of the browser client include handling movements of the participants, handling the shared events, handling the shared objects and providing communications with other participants. The VRML extensions include shared behaviors, objects and avatars. The pre-processor compiles extended parts of VRML. The wire protocol defines how the clients interact and exchange information about activities occurring in the virtual space, utilizing four protocols as follows: HTTP, TCP/IP, Best Effort Multicast IP and Reliable Multicast IP.

Carson is similar to the present invention only in the sense that Carson discloses controlling avatars. However, the teachings of Carson are distinct from that of the present invention in the following manner.

In Carson, action controls for the avatars (CG characters) and for the shared objects are not carried out by transmitting the behavior data to other clients. On the contrary, in accordance with the teachings of Carson, the behavior events specified by each client machine are notified to other client machines, and the notified behavior events are executed according to the behavior data stored in each of the client machines to reproduce the behavior events on a copy of the shared virtual space held by each client, thereby insuring the identity of the behavior events. In the machine which notifies the behavior events, the behavior events are executed according to the behavior data stored in the machine.

Carson fails to disclose behavior stream data. Accordingly, Carson fails to disclose changing behavior stream data. More particularly, similar to Matsuba, Carson does not disclose or suggest

correcting an input stream by replacing motion data of a selected component with data based on input operational contents in accordance with claims 38, 44 and 51-57. Accordingly, a combination of the teachings of Matsuba and Carson additionally fails to teach that which is required in each of independent claims 38, 44 and 51-57.

Naka teaches that a person can observe a CG animation in real time at a terminal. However, the user of terminal can only change the viewpoint at the terminal side and the user cannot control the CG characters and the objects in the environment. Further, Naka discloses, as opposed to Matsuba which discloses inputting motion information of CG objects, regulations of transmission data format of a stream. In particular, in accordance with the teachings of Naka, it is understood that the motion information of the CG objects are already possessed by a terminal user.

Naka fails to teach the shortcomings of Matsuba and Carson such that a combination of teachings of Matsuba, Carson and Naka would teach that which is required in each of independent claims 38, 44 and 51-57. In particular, similar to Matsuba and Carson as discussed above, Naka fails to teach replacing motion data of a selected component with data based on inputted operational contents.

Accordingly, a combination of the teachings of Matsuba, Carson and Naka fail to teach that which is required in independent claims 38, 44 and 51-57.

As described in paragraph 6 of the Office Action, Svancarek is cited for disclosing table conversion data to convert manual control data to motion data. It is respectfully submitted that Svancarek fails to teach the shortcomings of Matsuba, Carson and Naka such that a combination of the teachings of Matsuba, Carson, Naka and Svancarek would teach that which is required in independent claims 38, 44 and 51-57. In particular, similar to Matsuba, Carson and Naka as discussed above, Svancarek fails to teach replacing motion data of a selected component with data based on inputted operational contents.

As described in paragraph 7 of the Office Action, Bidiville is cited for disclosing a neural network to convert the movement of a trackball into X and Y components for movement of a cursor on a video display. It is respectfully submitted that Bidiville fails to teach the shortcomings of Matsuba, Carson and Naka such that a combination of the teachings of Matsuba, Carson, Naka and Bidiville would teach that which is required in independent claims 38, 44 and 51-57. In particular,

similar to Matsuba, Carson, Naka and Svancarek as discussed above, Bidiville fails to teach replacing motion data of a selected component with data based on inputted operational contents.

In light of the above discussion, it is respectfully submitted that claims 38, 44 and 51-57 would not have been obvious over the combination of the teachings of Matsuba, Carson, Naka, Svancarek and Bidiville within the meaning of 35 U.S.C. § 103. Furthermore, as claims 39-43 and 45-50 are dependent upon claims 38 and 44, respectively, it is additionally respectfully submitted that claims 39-43 and 45-50 would not have been obvious over the combination of Matsuba, Carson, Naka, Svancarek and Bidiville under 35 U.S.C. § 103.

Having fully and completely responded to the Office Action, Applicants submit that all of the claims are now in condition for allowance, an indication of which is respectfully solicited.

If there are any outstanding issues that might be resolved by an interview or an Examiner's amendment, the Examiner is requested to call Applicants' attorney at the telephone number shown below.

Respectfully submitted,

Yoshiyuki MOCHIZUKI et al.

By: Thomas D Robbins

Thomas D. Robbins
Registration No. 43,369
Attorney for Applicants

TDR/abm
Washington, D.C. 20006-1021
Telephone (202) 721-8200
Facsimile (202) 721-8250
May 13, 2003

VIRTUAL SPACE CONTROL DATA RECEIVING APPARATUS, VIRTUAL SPACE CONTROL DATA TRANSMISSION AND RECEPTION SYSTEM, VIRTUAL SPACE CONTROL DATA RECEIVING METHOD, AND VIRTUAL SPACE CONTROL DATA RECEIVING PROGRAM STORAGE MEDIA

FIELD OF THE INVENTION

The present invention relates to a virtual space control data receiving apparatus, a virtual space control data transmission and reception system, a virtual space control data receiving method, and a virtual space control data receiving program storage medium and, more particularly, to those for controlling a dynamic virtual space represented by three-dimensional computer graphics (hereinafter referred to as 3-dimensional CG), static image, dynamic image, audio, and text which are based on a network such as the Internet.

BACKGROUND OF THE INVENTION

10 In recent years, virtual malls, electronic commerce, and related home pages, such as WWW (World Wide Web) on the Internet, have attracted attention as utilization fields of 3-dimensional CG. Especially, the rapid progress of the Internet provides an environment in which relatively high definition 3-dimensional CG such as games and movies are easily handled at home. In the conventional WWW, a machine called a server, such as a personal computer or a work station, is connected through the Internet to plural machines called clients, such as personal computers. In this system, data such as video, audio, text, window layout, and

the like are downloaded from the server in response to a request from a client, and the client reconstructs the downloaded data to obtain necessary information. A communication method based on TCP/IP (Transmission Control Protocol/Internet Protocol) is employed for the server-to-client communication.

In the conventional WWW, data supplied from the server were mainly text data and video data. In recent years, with standardization of VRML (Virtual Reality Modeling Language) and browsers for VRML, there is a movement on foot to transmit 3-dimensional CG itself, such as shape data and texture data constituting a scene.

Hereinafter, the VRML will be briefly described.

In the conventional data format mainly composed of video data and text data, such as HTML (Hyper Text Markup Language), enormous time and cost are required for transmitting video data, especially, animation data. Therefore, in the existing system, network traffic is restricted. On the other hand, in the conventional 3-dimensional CG, all of data including shape data, view data, and luminous data are processed as 3-dimensional data. With the progress of 3-dimensional CG technology, the quality of created image is improved rapidly, and the efficiency is significantly improved with regard to the data quantity when 3-dimensional CG data is transmitted as it is. Usually, the data compression ratio in the case of transmitting 3-dimensional CG data is 1/100 or more as compared with the case of transmitting

represented by 3-dimensional CG, static image, dynamic image, audio, and text which are based on a network such as the Internet, the viewer can select an object or a part of an object to be controlled, and move it as he/she desires.

According to an eleventh aspect of the present invention, a virtual space control data receiving apparatus comprising: stream data receiving means for receiving stream data, and dividing the stream data into motion stream data and other stream data to be output; manual control data input means for inputting control data for an object or a part of an object to be motion-controlled manually; manual control data conversion means for converting the control data input by the manual control data input means, into motion data suited to the object or part to be controlled; and motion control data output means for outputting, as scene generation motion data, the motion data output from the manual control data conversion means, for the object or part to be controlled with the control data which is input by the manual control data input means, and outputting the motion stream data supplied from the stream data receiving means, for the other objects or parts. Therefore, in a dynamic virtual space represented by 3-dimensional CG, static image, dynamic image, audio, and text which are based on a network such as the Internet, the viewer can move objects or parts to be controlled, as he/she desires, by using the same control data.

According to a twelfth aspect of the present invention, the

means 5. Since the control data output means 4 is a kind of switcher, it is provided with a table describing identifiers of the respective control objects and information as to whether the respective control objects are based on the audio information or the scene information, and output data and their destinations are decided on the basis of the table.

Next, the synchronous operation of the control data output means 4 will be described with reference to figure 2. The overwrite buffer 9 receives the converted control data from the manual control data conversion means 3 during the frame playback period, writes the data while updating it, and outputs the ~~lately~~ recently written data. The audio/scene identification means 10 identifies the received stream data from the stream data receiving means 1, sends the audio information to the audio output means 8, and writes the scene information into the FIFO 11. On receipt of a synchronous signal, the synchronous output means 12 reads data from the overwrite buffer 9 and the FIFO 11, and outputs scene information. At this time, if the scene information in the FIFO 11 overlaps the converted control data written in the overwrite buffer 9, only the converted control data is output from the overwrite buffer 9 while the overlapping scene information is not output from the FIFO 11.

The scene data generation means 5 generates a scene at each frame time on the basis of the scene information transmitted from the control data output means 4 and the 3-dimensional CG data for

the stream data received by the stream data receiving means for the other objects. Therefore, in a dynamic virtual space represented by 3-dimensional CG, static image, dynamic image, audio and text which are based on a network such as the Internet, the motion of an object controlled by another virtual space control data receiving apparatus can be reproduced.

[Embodiment 3]

Figure 4 is a block diagram illustrating the structure of a virtual space control data transmission and reception system according to a third embodiment of the present invention. This system comprises a stream data transmission means 51, a client unit A 52, a client unit B 53, a manual control data transmission means 54, a data transmission/reception line 55, a stream data receiving means 56, a manual data input means ~~67~~ 57, a manual data transmission means 58, a manual data receiving means 59, a manual control data conversion means 60, a control data output means 61, a scene data generation means 62, a drawing means 63, a display means 64, and an audio output means 65. The structure of the client unit B 53 is identical to that of the client unit A 52. While in this third embodiment two client units are used to explain the processes performed by the virtual space control data transmission and reception system, the contents of the processes are identical even when three or more client units are used. Therefore, a virtual space control data transmission and reception system having three or more client units is also within

the scope of this third embodiment.

The scene data generation means 62, the drawing means 63, the display means 64, and the audio output means 65 are identical to the scene data generation means 5, the drawing means 6, the display means 7, and the audio output means 8 according to the first embodiment, respectively.

The respective constituents of the virtual space control data transmission and reception system so constructed will be described in detail.

The stream data transmission means 51 transmits stream data through the data transmission/reception line 55, like the stream data transmission means 21.

In the client unit A 52, the stream data receiving means 56 receives the stream data transmitted through the data transmission/reception line 55 and processes the stream data, in like manner as described for the stream data receiving means 25 of the second embodiment. The received stream data is output to the control data output means 61.

The manual data input means 57 outputs inputted selection data to the manual data transmission means 58, the manual control data conversion means 60, and the control data output means 61. Further, it outputs inputted control data to the manual data transmission means 58 and the manual data conversion means 60.

On receipt of the selection data and the control data output from the manual data input means 57, the manual data transmission

by directly solving a physical equation, the balance with the calculation time should be considered.

The motion control data output means 74 outputs, as scene generation motion data, the motion data supplied from the manual control data conversion means 73, for an object or a part of an object to be manually controlled, and outputs the motion stream data transmitted from the stream data receiving means 71, for the other objects or parts. In this case, amongst the objects or parts to be motion-controlled, those to be manually controlled are fixed or given identifiers.

The scene data generation means 75 generates scene data from the scene generation motion data at each frame time, which is output from the motion control data output means 74, and from other data required for scene configuration (e.g., 3-dimensional shape data, camera data, texture data, luminous data, data for bump mapping, data for illuminance mapping, etc.) which are externally supplied ~~from the outside~~. The scene generation motion data is motion data which is time series data by which the position of a moving object or the status of a skeletal structure at each time can be calculated. A transform sequence or the like is obtained from the motion data, and a 3-dimensional shape which defines the control object is transformed to the status of the 3-dimensional shape at each time (e.g., the positions of apexes of polygons constituting the 3-dimensional shape). Scene data is obtained by adding, to the motion data, other CG data indicating the shapes

of objects other than the target object, the status of camera, the texture mapping method, and the state of light source. That is, scene data is data required for generating a 3-dimensional CG image at each time.

The drawing means 76 generates a 3-dimensional CG image from the scene data output from the scene data generation means 75. As a 3-dimensional CG image generation method, Phong shading or Gouraud shading, which are generally known as luminance calculation methods, is used. As a hidden surface removal method, Z buffering or scan line buffering is used. Further, when using texture mapping, bump mapping, illuminance mapping, or shadow mapping, the reality is increased and thereby the image definition is improved. The image data of the 3-dimensional CG image generated by the drawing means 76 is displayed by the display means 77. A 3-dimensional CG drawing board on the market can be used as the drawing means 76, and a CRT or a liquid crystal display can be used as the display means 77.

The respective processes according to this fourth embodiment are performed in synchronization with each other. Especially, performing synchronization processing in the motion control data output means 74 is effective for pipelining the processes from generation of scene data to display of image data.

As described above, the virtual space control data receiving apparatus according to the fourth embodiment is provided with the stream data receiving means for receiving stream data and

image, audio and text which are based on a network such as the Internet, the operator is able to arbitrarily select an object or a part of an object to be controlled, and move the selected object as he/she desires.

[Embodiment 6]

Figure 7 is a block diagram illustrating the structure of a virtual space control data receiving apparatus according to a sixth embodiment of the present invention. The virtual space control data receiving apparatus comprises a stream data receiving means 91, a manual control data input means 92, a manual control data transmission means 93, a manual control data receiving means 94, a manual control data conversion means 95, a motion control data output means 96, a scene data generation means 97, a drawing means 98, and a display means 99.

The stream data receiving means 91, the scene data generation means 97, the drawing means 98, and the display means 99 are identical to the stream data receiving means 71, the scene data generation means 75, the drawing means 76, and the display means 77 according to the fourth embodiment.

Hereinafter, the respective constituents of the apparatus will be described in detail.

22 The manual control data input means 92 is used for inputting control data or motion data like the manual control data input means 72 of the fourth embodiment, and sends the inputted control data or motion data to the manual control data transmission means

93 and the manual control data conversion means 95 ~~94~~.

The manual control data transmission means 93 transmits the control data or motion data from the manual control data input means 92 to an external virtual space control data receiving apparatus which has the same structure as that of this sixth embodiment. On the other hand, the manual control data receiving means 94 receives control data or motion data transmitted from the external virtual space control data receiving apparatus, and outputs it to the manual control data conversion means 95.

Hereinafter, the method of transmitting and receiving control data will be described by using figures 13(a) and 13(b). Figure 13(a) shows the format of a control data packet corresponding to one block, and transmission and reception of control data are performed using this packet. The header section of the control data packet comprises client identifiers given to a plurality of virtual space control data receiving apparatuses, a packet identifier indicating that this packet is a control data packet, a time stamp indicating a time from a reference point of time at which this packet was generated, and the total number of channels (Dc) to be transmitted. The data section comprises, for one channel, a channel identifier indicating a channel number, and compressed or non-compressed data (data to be transmitted) equivalent to the packet size. That is, the data section comprises the channel identifiers and the data to be transmitted as many as the number of channels (Dc). As shown in figure 13(b),